TM-(L)-7[illegible]/00/01

TECHNICAL MEMORANDUM

(TM Series)

This document was produced by SDC in performance of contract Government Contracts

INFORMATION MANAGEMENT--A STATE-OF-THE-ART STUDY

Part I - Synopsis of Organizations and Research Activities

Carole E. Bare

8 October 1962

SYSTEM DEVELOPMENT CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA

ABSTRACT

Information retrieval, which has traditionally included indexing and abstracting activities, is seen as only one of the major functions in an information system. A more broadly conceptualized information system should provide control over more varied and complex functions, such as those involved in problem-solving. Therefore, research activities in this area, which have up to now been labelled as information retrieval studies, will be discussed in this report from the more comprehensive viewpoint of an "information management" system.

A state-of-the-art study in the information management area was undertaken since participation in this rapidly growing field should be based on sound decisions regarding the content areas as well as the methods and techniques which are to be applied. In this first report, a synopsis of research organizations and their sponsors is presented. In addition, several classification systems of the various functions of an information management system are described and evaluated. Trends which indicate the direction in which the field is moving are outlined.

TABLE OF CONTENTS

INTRODUCTION

It has become increasingly clear that the presently used concept of "information retrieval" is no longer adequate to explain the full scope of the area. All too frequently "information retrieval" has become synonymous with document storage and retrieval and the most attention, therefore, has been centered around the mechanical and data processing aspects of these library type functions. To satisfy a wide variety of criteria (for instance, information requirements for efficient decision making) a more comprehensive conceptualization must be provided which gives proper emphasis to:

1. User information requirements as a basis for selection of data for input to the information system.
2. Computer capability to reorganize and manipulate the stored facts.
3. Closing of the linguistic gaps (human vs. computer language) between man and computer in relation to the formulation of requests.

In addition to performing retrieval functions, a system based on these broader concepts would provide control over more varied functions, e.g., those involved in problem-solving, and, therefore, would more appropriately be called an "information management"* system. An information management system differs from an information retrieval system in two major ways: (1) The information retrieval system includes only a limited number of functions, and the major accent is placed on the search for a previously stored document or fact, whereas the information management system encompasses the whole spectrum of information handling functions: How the information which the system handles is used determines its selection, its format for input, storage and inquiry as well as the search strategies employed. (2) In addition, the kind of functions which an information management system is able to handle are more complex. Again depending upon the use to which the information is put, the system will be able to have

*Term introduced in "Information Management - Long Range Planning Report," Stanford Research Institute, Menlo Park, California - Report 39, February 1960.

reorganization features such as correlation of facts, automatic modification of its memory system (updating and purging, often called retrospective indexing) and its executive program according to pre-established rules.

Although retrospective indexing research has recently been conducted under the label of information retrieval, it really represents a more dynamic and complex function. The growing interest in this research attests to the need for a reconceptualization. The more inclusive term will be used here to refer to this more comprehensively conceptualized field.

PURPOSE

Research on information management systems has mushroomed during the past decade. The major factors which are involved with this phenomenal growth rate have to do with aspects of the growing complexity of civilization:

- growing volume of technological and scientific information;
- increasing shortage of technical manpower;
- duplication of scientific and technological efforts;
- increased cost of research;
- increasing awareness of man's memory limitations with respect to pertinent information for problem solving;
- need for rapidly accessible, accurate, and up-to-date information for decision making and exchange of ideas and facts.

Most of the above factors represent the principal goals toward which research efforts in the information management field are directed. These goals will serve as a basis for a criterion framework which will have to be considered for the evaluation of research activities.

To engage in research and development on specific aspects of information management, it is not enough to know something about the content matter of the particular area under exploration; one must also become familiar with the methods and techniques being used and developed throughout the entire rapidly

growing field. Current research trends need to be analyzed so that the directions in which the field is to move can be consciously chosen. To satisfy these information needs an ongoing state-of-the-art study has been started.

After a preliminary exploration of the nature of the information management problems, the first step was to assess the make-up of the community within which research programs are being developed. Further steps will include a more intensive and continuously updated picture of content, locations, size, quality, and sponsorship of research efforts, as well as the identification of problem areas and trends As information on additional programs and new facets is added, the initial framework developed for studying this field may have to be reorganized to fit changing ideas and new technology; basic criteria such as rigor, economy, simplicity, adaptability, and compactness used to evaluate the research may also have to be identified. This reorganization of the conceptual framework and the modification and re-weighting of the criteria will be part of the ongoing assessment of the state-of-the-art.

METHODS

The methods employed in conducting this ongoing study include the collection of written or spoken information about ongoing or planned efforts, participation in symposia and conventions of the information management research community, and the development and application of the criterion framework for evaluating the quality and the trends of the research efforts.

In analyzing the changing research picture, the most difficult task is the assessment of the size of current efforts. Many organizations are hesitant to give out specific information concerning numbers of personnel and dollars expended on particular projects. In addition, the overhead figures are not computed on a comparable basis and the ratio of professional to support staff is not readily available. Frequently, efforts in this field are deeply embedded in the context of projects labeled simulation and modeling, man-machine integration, operations research, and systems analysis and design; thus making

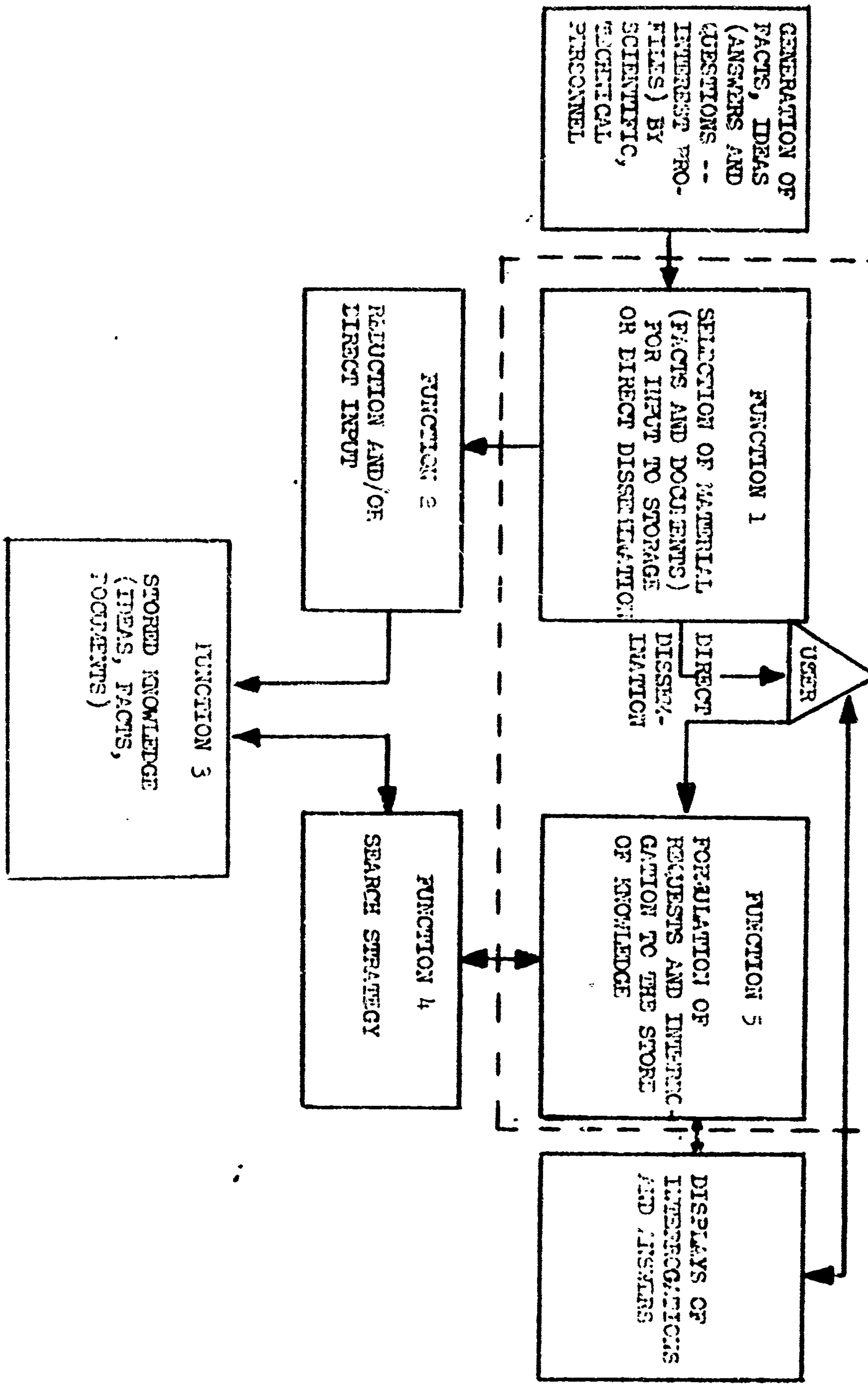
the task of identifying costs of research and development pertinent to this study area an additionally difficult one. Nevertheless, some useful information is available and can be translated into report format, although this information cannot presently be stated in a very precise manner. Hopefully, the volume of available information will increase as the amount of research and development work in this area increases and as SDC personnel become better acquainted with sponsors and researchers who have vested interests in the information management field.

ORGANIZATION OF RESEARCH ACTIVITIES

In discussing research activities, there is the problem of how to meaningfully categorize the work that is being done in the information management field. As mentioned in the introduction, a broader framework than has been used in the past needs to be developed for this purpose. Such a framework should take into consideration the five major information management functions shown in Figure 1.

Figure 1 contains a simplified diagram describing the flow of information management functions in an information management model. The first two functions are selection (1) and reduction (2) of information. Information which is eventually used for problem-solving, decision-making, research, etc. is selected for input to the store from the extensive reservoir of ideas and facts generated by scientific, technical and professional personnel. This information is either stored in natural language or reduced to a special language format by indexing or abstracting techniques. The next three functions concern storage (3), search strategy (4), and the formulation of requests (5). The requester may ask the computer to produce a list of documents or facts containing information he needs for problem-solving or he may want the computer to manipulate the stored data so that relationships of the data may bring to light information which was not previously known. Such requests may be formulated by several methods: the requester may be able to make inquiries to the computer in natural language and, through several conversations (teaching machine or other methods), the questions may become more precise; or he may have to add to his own vocabulary

FIGURE 1. INFORMATION MANAGEMENT SYSTEM
SIMPLIFIED DIAGRAM ILLUSTRATING FUNCTIONS AND INFORMATION FLOW
USER INFORMATION REQUIREMENTS AND NEEDS



the language of the information management system. A disadvantage of the latter method is that the searcher is forced to think like the indexer and thus the degree of creativity of searching techniques is limited to parameters foreseen by the indexer. A combination of these two methods may hold the best promise for optimum man-machine interaction; thus the computer can at increasingly sophisticated levels learn to duplicate intellectual processes leading to discovery as it browses through a broad range of data with purpose and the capacity to correlate facts.

The ways in which the above functions are fulfilled are determined by (1) the user needs and (2) the technological (equipment and system design) feasibility of implementing the methods which are to implement these user needs. The output products which the information management system provides for the user may be retrieved documents and facts, or new concepts, which the computer found and compiled by manipulating configurations of facts contained in its store.

It is difficult to divide the input/output cycle of an information management model into distinctly separate procedures or subsystems since most of the functions are interdependent. However, to discuss present research activities and to decide the direction of future efforts, the information management system must be analyzed into meaningful parts. For this purpose, the categories which have served as a traditional reporting framework during the years since 1948 will be modified and reorganized. This is only an interim step leading eventually to the development of a set of categories which is more in keeping with the description of the five functions mentioned above. It is not feasible to go to this future schema now because current research and development work in this field is not usually reported in these terms. It is hoped, however, that this new schema will not only provide a more meaningful report framework, but will also aid in defining more clearly the direction which research and development work in this field should take.

The following discussion explains why and how the four interim categories were developed from the categories into which research has been classified previously.

These four interim categories will hopefully provide a meaningful bridge to a proposed future schema which will be discussed later in this report. This future schema will treat the five functions as the basic categories; and for each of these functions, considerations of user needs and equipment research become requirements and constraints to the information management model.

The National Science Foundation, which has been the pioneer in the field, has been publishing biannual summaries of "current" research and development activities since 1948; in their reports, the research has been classified into the following five categories:

1. Information Needs and Uses
2. Information Storage and Retrieval
3. Mechanical Translation
4. Equipment
5. Potentially Related Research

The disproportionate growth of category 5 shows that the scope of the area is widening, and by now this collection of studies has all the earmarks of a "miscellaneous" category. When the studies included in this category are sorted out, it becomes evident that a redefinition of several other categories has become overdue. For instance, three of the areas subsumed under "Potentially Related Research" - Character and Pattern Recognition, Speech Analysis, and Linguistic and Lexico-graphic Research - can be classified as indexing and abstracting processes. They can also be subsumed under the Formulation of Requests and are closely connected with Search Strategies (see Figure 1.). These topics as well as the research in Artificial Intelligence and those categorized as Psychological Studies (also listed under "Potentially Related Research") are no longer potentially related. They have become a very core of the research field.

On the basis of current research and development activities, a regrouping of functions is urgently needed at this time. Furthermore, due to the "in flux" nature of this new field, the interrelationship of the above topics should be periodically re-evaluated. Some discernable trends which indicate that a change

in conceptualization of the system is called for also indicate cost directions which the information management research will take in the future. It is, therefore, necessary to provide a conceptual framework with the flexibility necessary for updating the emerging schema. The predictions for future direction and a preview of the next framework will be discussed and documented more fully in the section on trends. For the present synopsis, the research activities in the information management area are divided into four major categories:

1. User Requirements and Needs (including selection of information)
2. Reduction (indexing, abstracting and storage)
3. Search Strategy (including formulation of requests, translation)
4. Equipment

The first category, "User Requirements and Needs," includes studies (such as surveys) on user groups with respect to their needs for formally or informally obtainable information, publication studies (kind and quantity of published material with or without user evaluation of its usefulness), studies on communication processes, etc. Many of these studies serve to clarify the criteria for selecting ideas, hypotheses, questions, and facts to be entered into the information store.

The second category, "Reduction," includes studies covering the data reduction process through which selected material is prepared for input to the computer either by manual or machine abstracting or indexing techniques.

The third category, "Search Strategy," includes studies on search strategy as well as linguistics, parts of artificial intelligence and mechanical translation. Mechanical translation may also apply to the "Reduction" category. It was included here because the concept of search strategy is defined to be broader than just "retrieval." The computer not only searches and retrieves facts and documents but, by manipulating the data, performs a more sophisticated function: it may eventually be able not only to update but to change its own stored data configurations. Most of the linguistic research has a direct application to the formulation of the requests for information and to the ways in which the computer searches for it or correlates the existing information to

answer the question of the user.

The fourth category, "Equipment," includes at this time a gamut of studies on equipment. Although these studies emphasize the hardware more than the software approach, the strictly hardware approach to information management problems is generally on the decline. Equipment research is being conducted more in the context of broader information management problems. However, the number of studies showing the development and refinement of very specific equipment is still sizeable, requiring a separate category at this time.

The following four tables represent a brief synopsis of the agencies, names of contact people, available at this time, research topics, sponsorship, and estimates of the size of the research effort.

Table 1. USERS NEEDS - SELECTION OF INFORMATION FOR INPUT

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort*, etc.)
American Psychological Assoc. William D. Garvey Belver C. Griffith	Steps scientists take in recording and publishing scientific treatises - use of research made by the discipline - do papers or talks appear in topical summaries and formal literature?	National Science Foundation - small
Arthur D. Little, Inc. Jerome Herniter	System analysis of potential U.S. Science information system - experimentation with varying degrees of centralization of information - find optimum degree for dissemination purposes.	Medium
Bolt, Baranek & Newman J.L. Licklider - now at ARPA E. Fredkin	Man-Machine Communications	Council on Library Resources - medium
Case Institute of Technology Russell Ackoff	Measurement of value of recorded information.	National Science Foundation - small
Herner and Company Saul and Mary Herner	a. Determination of Usefulness of the National Institute of Health/National Library of Medicine Translations from Russian.	National Library of Medicine - small
	b. Factors which influence announcement, availability, and publication of scientific information.	National Science Foundation - small
Institute for Advancement of Medical Communication Richard Orr	a. Pre-publication dissemination study	U.S. Department of Health, Education and Welfare - small
	b. Selection procedures for input - consistency of ratings studied.	

* Small = estimated to be under $500,000; Medium = estimated to be from $500,000 to $1,000,000; and Large = estimated to be over $1,000,000.

Table 1. USERS NEEDS - SELECTION OF INFORMATION FOR INPUT (CONTINUED)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort, etc.)
John I. Thompson, Inc.	Study of dissemination of federally sponsored research - 12 Universities and Libraries across the country serve as regional technical report centers - each will make available reports by 3 federal agencies (90%) DOD - AEC - NASA.	Small
National Bureau of Standards Ethel Marden	a. System analysis of Office of Technical Services - feasibility of establishing Department of Commerce as clearing house for technical information.	various sponsors - medium
	b. Needs analysis for Food and Drug Administration	various sponsors - medium
Stanford Research Institute Jerre D. Noe	"Augmented Human Intellect Study" - man-machine relation between problem-solving and automated aid.	Self and U.S. Air Office of Scientific Research - small
System Development Corporation R. Harrington	a. Automated information processing for medical information systems.	Medium
	b. Development of prototype support center for medical research.	
	c. Research on patient data automation.	
United States Patent Office D. D. Andrews E. Glazer Dick Spencer	Preliminary determination of requirements for machine retrieval of case law and patent law.	Small

Table 1. USERS NEEDS - SELECTION OF INFORMATION FOR INPUT (CONTINUED)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort, etc.)
University of Illinois Charles E. Osgood	Structure of network of psychological journals - cross-citations - pattern of information exchange - considerations which psychologists have for their journals.	National Science Foundation - small
University of Michigan, Mental Health Research Institute - Richard L. Meier	Information input overload in large library.	Small

Table 2. REDUCTION OF DATA FOR INPUT AND STORAGE (EXCLUDING LINGUISTIC RESEARCH)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort*, etc.)
American Institute for Research, Pittsburgh	Guidelines for indexing.	Small
Datatrol Corporation Hugh P. Donahue William Hammond Arthur D. Kirsch	Comparison of ASTIA-AEC descriptors with the purpose to find compilers (equivalents).	Small
Documentation, Inc. Mortimer Taube	Experimental report to the Atomic Energy Commission on nuclear science abstracts - abstract and citation journal with machine producing index.	NASA, USAF (Rome) - large
General Electric National Library of Medicine Seymour I. Taine	Medical literature analysis and retrieval system (MEDLARS) - compilation of bibliographies.	National Library of Medicine - large
Herner and Company Saul and Mary Herner	Index format - telegraphic abstracts - cross indexing.	USAF, Office of Scientific Research - small
IBM Advanced Systems Development Division, San Jose Marjorie Griffin	Bibliographies - permutation indexes on 704 - Thesaurus building - generally, mechanization of library routines.	Self - small to medium
National Bureau of Standards	Abstracting of articles for broad and specialized classification with Western Reserve System - GE 225.	National Institute of Health and American Diabetics Association, University of Minnesota - large

* Small = estimated to be under $500,000; Medium = estimated to be from $500,000 to $1,000,000; Large = estimated to be over $1,000,000.

Table 2. REDUCTION OF DATA FOR INPUT AND STORAGE (EXCLUDING LINGUISTIC RESEARCH) (CONTINUED)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort, etc.)
Ramo-Wooldridge, Division of Thompson Ramo Wooldridge, Inc. Harold P. Edmundson Paul L. Garvin J. L. Kuhns L. C. Ray	a. Automatic indexing.	Council on Library Research, Inc. - small
	b. Inductive methods in semantic analysis with the goal of developing reliability in compiling of semantic codes for automatic abstracting, etc.	USAF, Office of Scientific Research - small
System Development Corporation Harold Borko Lauren Doyle	a. Evaluation of automatic indexes.	Small
	b. Research on statistical properties of text.	
	c. Linguistic feedback.	
U. S. Patent Office - Research and Development D. D. Andrews	PACIR project - semi-automated coding techniques for chemical compounds (insecticide) - information usable in several different machines.	Small to medium

Table 3. SEARCH STRATEGY (INCLUDING LINGUISTIC RESEARCH)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort*, etc.)
Advanced Information Systems Co.	Self-organizing file with adaptive capability.	Small
American Legal Data Processing Association	Voice desk retrieval in conjunction with microfilm and Univac 490.	Ryan Electric; Transdata; Univac, Sperry Rand - small
Bell Telephone Laboratories, Inc. L. D. Harmon	Computer reading of cursive script.	Self - small
Computer Associates	Retrieval requests - monitoring system for file reorganization.	Self - small
Douglas Aircraft Company Jim Morrison L. R. Bunnow	Started with mechanization of file - then used interest profiles - relates to pre-publication information dissemination - use SDI - see IBM	Self - small
Harvard University Philip J. Stone	Automated content analysis scheme.	National Institute of Mental Health - small
IBM - Advanced Systems Development Division, Yorktown Heights R. E. Nienburg T. R. Savage	Selective Dissemination of Information (SDI) - coding and programming languages	Self - small
IBM - Thomas J. Watson Center	Concept-processing: investigate possibility of machine input centered about a special kind of logical memory organization capable of supplying names or sentences appropriate to certain requests.	Self - small

* Small = estimated to be under $500,000; Medium = estimated to be from $500,000 to $1,000,000; Large = estimated to be over $1,000,000.

Table 3. SEARCH STRATEGY (INCLUDING LINGUISTIC RESEARCH) (CONTINUED)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort, etc.)
ITEK Corporation R. F. Barnes L. F. Buckland E. A. Lipetz J. W. Kuipers	Information retrieval system built by grammatic normalization to get subject index data.	Small
Massachusetts Institute of Technology Victor Yngve	Linguistic research leading toward mechanical translation of language.	National Science Foundation - small
North American Aviation, Inc. (Autonetics) J. D. Bledsoe	Pattern recognition - waveform data and image identification - speech	Self - small
Stanford Research Institute Jerre D. Noe	Simultaneous interrogations of data store with multiple responses.	USAF (Rome) - small
System Development Corporation a. R. Simmons	a. Synthex - Computerized systems for synthesizing complex human cognitive functions.	Self - small
b. Aiko M. Horman L. E. Travis S. S. Shaffer	b. Complex learning by machine.	Self - small
c. Charles Vossler Leonard Uhr	c. Simulation of brain model.	Self - small
Thompson Ramo-Wooldridge, Inc. H. P. Edmundson P. L. Garvin J. L. Kuhns L. C. Ray	Automatic abstracting of English language text to facilitate machine searches.	USAF (Rome) - small

Table 3. SEARCH STRATEGY (INCLUDING LINGUISTIC RESEARCH) (CONTINUED)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort, etc.)
Western Reserve University J. H. Shera Allen Kent	"Current awareness" and "retrospective search" varies search at several levels of detail.	NSF - small

Table 4. EQUIPMENT (INCLUDING COMPUTER TECHNOLOGY AND DISPLAY)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort*, etc.
General Electric, Schenectady W. E. Glenn	Reliability of electronic components - also thermoplastic recording.	Self - small
Herner and Company H. T. Heatwole Saul Herner	Electronic information and data searching and correlating devices.	Self - small
IBM, San Jose Norman Vogel	Walnut - large capacity random access document retrieval.	Self - medium
John Diebold Associates D. Blumberg	All phases of equipment research - state-of-the-art report on equipment for information management.	Medium
Massachusetts Institute of Technology D. M. Baumann	Photo memory - photo storage techniques.	Office of Naval Research - small
MITRE Corporation E. M. Bennett R. P. Mayer P. R. Bagley	a. Demonstration of computer capability to accept input of natural language (8-year old level). b. Acceptance by computer of military requests (commands in basic English and to take action in a simulated environment).	Small
National Bureau of Standards a. S. N. Alexander b. T. C. Bagg J. L. Pike	a. PILOT - high speed electronic digital data processing system. b. RAPID SELECTOR - also input devices.	a. Self - small b. Bureau of Ships, Dept. of Navy - small

* Small = estimated to be under $500,000; Medium = estimated to be from $500,000 to $1,000,000; Large = estimated to be over $1,000,000.

Table 4. EQUIPMENT (INCLUDING COMPUTER TECHNOLOGY AND DISPLAY) (CONTINUED)

ORGANIZATION Contact Persons	TOPIC OF RESEARCH AND DEVELOPMENT	COMMENTS (Sponsorship, Size of Effort, etc.)
National Bureau of Standards (cont.) c. Joshua Stern	c. Oscillograph paper for more rapid print-out - modification of the solenoid-driven punch - utility and improvement of the diazo card in peek-a-boo deck.	c. Self - small
National Cash Register C. O. Carlson D. A. Grafton A. S. Tauber	Photochromic micro image - storage and retrieval of document images.	Self - small
Radio Corporation of America Astro-Electronic Division M. S. Cohen Jack Mincker	ACSI-MATIC - intelligence data system - development of inquiry console: intelligence analyst will be able to communicate his requests directly to the machine in simplified English and standard military terminology.	Self - medium
Stanford Research Institute James D. Noe	Transfer of graphic data on maps, charts and aerial photographs to storage for later retrieval.	U. S. Army Signal Corps - small

CHARACTERISTICS OF SPONSORSHIP OF RESEARCH

Table 5 presents a synopsis of what is currently known about the major organizations which sponsor research in the information management field. Several of these organizations also support in-house research activities in this area. However, this analysis includes only their support of research efforts outside the agency. In addition to the sponsors listed in Table 5, there are numerous sponsors who were not listed and who are supporting only one outside effort on a budget less than $50,000 per year. The fact that many minor sponsors are involved in small research projects emphasizes not only the growing interest in this type of research but also a certain degree of fractionation.

The most important characteristic of the funding of research activities in information management is that of joint sponsorship. This fact has far-reaching implications for organizations who are planning a major involvement in the field.

The estimates of the size of support did not vary by sponsor from 1961 to what is known at this time about 1962. This means that the funding remained relatively constant within the small, medium, and large financial support categories. The number of studies, however, has increased for several sponsoring agencies, implying a decrease of support funds per study. The aggregate sum of support in no case was estimated to exceed $2.5 million; many of the support estimates hover around the million dollar figure. The USAF and U.S. Army seem to be providing the largest support for information management research.

MAJOR SPONSORS OF RESEARCH ACTIVITIES IN INFORMATION MANAGEMENT*

Table 5

Name of Sponsor	Estimated Size of Support Per Year - Latest Estimate			Number of Studies Sponsored	
	Large	Med.	Small	1961	1962
Arthur D. Little, Inc.			X	-	1
Autonetics			X	-	1
Carnegie Institute of Technology (together with RAND Corp. and University of Calif.)			X	2	2
Council on Library Resources (Ford Foundation, budget eight million)	X			12	12
EURATOM (Europe) (together with European Institutes)		X		7	5
IBM		X		1	2
Japanese Ministry of Education			X	2	2
Lincoln Laboratory (MIT)			X	1	1
[illegible]			X	1	1
NASA	X			2	2
National Institutes of Health Unspecified Nat'l Heart Institute Nat'l Inst. Mental Health (jointly with several other sponsors, such as NSF, U.S. Atomic Energy Comm., USAF (OSR), etc.)	X	X	X	7 - -	11 1 3

Large - From $500,000 - $2,500,000
Medium - From $ 50,000 - $ 500,000
Small - Under $ 50,000

* Sources of Data: National Science Foundation, other descriptions of research programs and conversations with resource personnel in the field.

Table 5 (Continued)

Name of Sponsor	Estimated Size of Support Per Year - Latest Estimate			Number of Studies Sponsored	
	Large	Med.	Small	1961	1962
National Science Foundation (jointly with many other sponsors like NIH, Universities, etc.)	X (Over 1 million in 1962)			37	57
Sperry Rand-Univac		X		-	1
U.S. Air Force (jointly with other sponsors)					
Unspecified		X		3	3
Aeronautical Systems Div. (Microwave Tech. Section)		X		1	5
Aeronautical Systems Command			X	-	1
Air Materiel Command			X	1	-
ARDC			X	-	1
Cambridge Research Lab.	X			3	3
Office of Scientific Res.	X			18	25
The RAND Corporation			X	2	-
Rome Air Development Center	X			16	35
Wright Air Development Division		X		2	1
U.S. Army (jointly with other sponsors)					
Unspecified (together with Air Force, Navy, Weather Bureau & NSF)		X		2	1
European Research Office			X	2	-
OACSI-(Office of Asst. Chief of Intelligence) (ACSI-matic - RCA)	X			1	1
Office of Ordnance Research (jointly with U.S. Signal Corps)		X		1	-
Office of the Surgeon General (University of Maryland)			X	1	1
Research Office		X		-	1
Signal Corps		X		2	3
Signal Research & Development Lab.		X		1	2

Large - From $500,000 - $2,500,000
Medium - From $ 50,000 - $ 500,000
Small - Under $ 50,000

Table 5 (Continued)

Name of Sponsor	Estimated Size of Support Per Year - Latest Estimate			Number of Studies Sponsored	
	Large	Med.	Small	1961	1962
Atomic Energy Commission			X	2	2
U.S. Department of Commerce					
National Bureau of Standards		X		1	2
U.S. Patent Office (jointly with NSF and Nat'l Bureau of Standards)		X		1	1
U.S. Weather Bureau (jointly with USAF, Army, Navy and NSF)			X	-	1
U.S. Navy (jointly with other sponsors)					
Unspecified			X	1	1
Bureau of Ships			X	1	1
Office of Naval Research	X			17	18
U.S. Office of Education (jointly with West Res and NSF)			X	1	1
U.S. Public Health Service		X		-	1

Large - From $500,000 - $2,500,000
Medium - From $ 50,000 - $ 500,000
Small - Under $ 50,000

TRENDS

Before discussing the major, basic trends which emerge from this study and their implications, several problem areas which are receiving increased attention and emphasis should be mentioned. The publication lag of technical and scientific information is a problem which may be alleviated through information management research; this lag also becomes a serious stumbling-block to efficient and timely utilization of the findings of research activities in this fast-moving field. As a result, much duplication of effort occurs. In addition, due to the relative newness and fluidity of the field, the quality of research reports is poor. Information about the general purposes, specific criteria, the methodology and findings is seldom clearly and explicitly stated. These difficulties are due mainly to the larger number of exploratory rather than research studies. However, work is in progress at least to change the report format of studies to provide more precise information.

Another emphasis which is worth mentioning is the increased evidence of interest in information management in countries other than the United States. However, there seems to be much duplication of effort between the foreign and American research. If American investigators are able to disseminate published results of research, update them and explain them clearly, other countries may be able to use the work as a starting point rather than having to begin from scratch.

Three major inter-related trends are discernible and should give direction to the planning for participation in the information management area:

- A definite shift from hardware to software orientation;
- An increase in amount, quality and depth of user requirement research;
- Some changes in the conceptualization of the information management field which may lead to a shift from component orientation to a system approach.

Only about 10% of the total studies in information management are devoted to equipment research per se. As has been true of much system design and research in the past, experimentation with hardware problems took place before the man-

machine roles were adequately defined. At times system designers seem to imply that it may be easier to retro-fit the human performance than that of the equipment. It is a hopeful sign for the development of this new field that there is a trend to design and test equipment as an adjunct to, or in the context of, storage, retrieval, request formulation, and display problems.

Thus, companies which have been primarily interested in the hardware approach are gradually beginning to subordinate these efforts to software considerations; and to some extent organizations such as IBM and G.E. are becoming involved in approaches which demand a more comprehensive conceptualization of the information management area. However, there does not seem to be much evidence of coordination among parallel research programs which are reported by the same companies in various problem areas. Geographical separation of the locations where the research is done may account for some of this; but a lack of coordination seems to be evident even when members of research teams conducting these various studies are located in the same division or location of the large organization. This fractionation may also be caused by the fact that when very specific technical information management problems occur, the organizations involved may have special capabilities and skills to solve these somewhat discrete problems; which, to their researchers, have the appeal of representing basic research in an area of high specialization.

The variety, quality and amount of user requirements studies has definitely increased. Studies are not confined any longer to the survey type research which had as its sole purpose to find out what information the scientists need, how fast they need it, and what they would be willing to pay for this service. Researchers in the user needs area are beginning to see as their domain the evaluation of the total status of the more broadly defined scientific communication process. The studies range from the traditional survey type (The American Institute of Biological Sciences) through explorations on the most efficient degree of centralization versus decentralization of stored information (Arthur D. Little, Inc.) to studies concerning the steps which scientists take sequentially while they discover and formulate new relationships of data (American Psychological Assn.).

This last-mentioned type of user needs study may lead to a clarification of the nature and formulation of the questions which the researcher and/or decision-maker has to put to the information store. The decision whether the subject matter is to be recorded at all (selection for input), whether it is to be recorded and retrieved in the form of total text in natural language, or as an abstract, or through an index of key words, unit terms, etc., depends mainly on the type of service to be provided. In addition, the method for organizing the subject matter is governed to a great extent by its nature, i.e., does it cover an established set of facts or ideas, hypotheses and notions, or a mixture of these?

It thus becomes clear how closely the different functions such as the selection of material for input to the store, the storage organization, the search strategy, and the request formulation are tied to the exploration of different depth probes into the User Needs problem area.

Because user requirements influence all phases of information management research, they must be viewed as the criterion framework for each of the five functions discussed earlier, i.e., selection for input, reduction, storage, search strategy and request formulation (display format can be included here). The explorations of problem areas and methodology should meet specific criteria made explicit by user requirement research. Equipment research will probably also have to be selected and reported on for each of these functions. Therefore, a type of schema can be proposed which should reflect desirable and growing trends in the conceptualization of the field. Figure 2 presents a projection toward this next step of the state-of-the-art study. If the pulse of change is felt correctly, user requirements and needs will be the guide to research activities for central problem areas (examples listed) with respect to the five functions. Equipment research will also occur and be reported in connection with these functions.

FIGURE 2. HYPOTHETICAL SCHEME FOR CHANGE IN CONFIGURATION OF FUNCTIONS IN THE INFORMATION MANAGEMENT FIELD

CRITERIA	USER HUMAN AID REQUIREMENTS -- documents, facts, ideas, hypotheses, updated and correlated fact configurations, direct disseminations from interest profiles, costs, etc.				
FUNCTIONS	1 SELECTION FOR INPUT	2 VARIOUS DEGREES OF INPUT REDUCTION	3 STORAGE	4 SEARCH STRATEGY	5 REQUEST FORMULATION - DISPLAY
EXAMPLES OF PROBLEM AREAS	Can computer select materials from interest profiles or from other criteria?	Human vs. computer abstracting and indexing? ---Manual vs. reduction?	What are good criteria for purging?	Fact vs. document search. To what extent can computer aid with, or make, decisions and do research?	Should computer learn requesters' language or vice versa, or some of each?
AVAILABLE EQUIPMENT	List of Equipment	List of Equipment	List of Equipment	List of Equipment	List of Equipment